AMENDMENTS TO THE **C**LAIMS

This listing of claims will replace all prior versions of claims in the application:

Listing of Claims:

1.   (Currently Amended) A system that refines a general-purpose search engine, comprising:

a component that identifies an entry point that includes a link utilized to access the general-purpose search engine; and

a tuning component that receives search query results of the general-purpose search engine and filters the search results based at least on criteria associated with the entry point through which the general-purpose search engine was accessed, the criteria comprises at least a first set of data categorized as relevant to a user's context and a second set of data categorized as non-relevant to the user's context, ~~each search result is compared with both the first set of data and the second set of data to determine a relevance of the search result~~[[,]] <u>wherein user selection of a query result from a ranked list of the query results causes the selected result to be added to the first set of data and causes the results not selected by the user but ranked higher than the selected result to be automatically added to the second set of data, the first and second sets of data persisted to a computer-readable storage medium.</u>

2.   (Original) The system of claim 1, the criteria comprising one or more of a document property, a context parameter, and a configuration.

3.   (Original) The system of claim 2, the document property comprising one or more of a term that appears on a web page, a property of a Uniform Resource Locator (URL) identifying the web page, a property of a plurality of URLs that link to the web page, a property of a plurality of web pages that link to the web page, and a layout.

4.   (Original) The system of claim 2, the context parameter comprising one of a word probability and a probability distribution

5.   (Original) The system of claim 1, the tuning component is provided with training data to learn what properties of a document are indicative of the document being relevant to a user executing a search query from the entry point.

6.   (Original) The system of claim 1, the tuning component configured to differentiate between a query result that is relevant to a search query context for a group of users and a query result that is non-relevant to the search query context for the group of users.

7.   (Previously Presented) The system of claim 1, the tuning component employs statistical analysis in connection with filtering the search query results.

8.   (Previously Presented) The system of claim 1, the tuning component generates one or more context parameters for a received query result, and compares the generated context parameters with a relevant context parameter and a non-relevant context parameter to determine whether the query result is relevant.

9.   (Previously Presented) The system of claim 1, the tuning component further ranks the query results.

10.   (Previously Presented) The system of claim 9, the ranking determined by the degree of relevance of the query result to the relevant data set and the non-relevant data set, the relevance is determined *via* one of a similarity measure and a confidence interval.

11.   (Original) The system of claim 9, the ranking order comprising one of ascending and descending, from the most relevant result to the least relevant result.

12.   (Original) The system of claim 1, the tuning component configured for a plurality of entry points associated with one or more groups of users.

13.   (Currently Amended) A system that tunes a general-purpose search engine, comprising:

a filter component that receives search query results of a general-purpose search engine and parses relevant and non-relevant results based on training data associated with the entry point that provides a link employed to traverse to the general-purpose search engine, the training data comprises a first set of data categorized as relevant to a search context of a user for the entry point and a second set of data categorized as non-relevant to the search context of the user, ~~each search query result is compared with the first set of data to determine a degree of relevance of the search result, and each search result is compared with the second set of data to determine a degree of non-relevance of the search result~~; and

a ranking component that sorts the filtered results in accordance with the training data for presentation to a user, <u>wherein a user clicking a link associated with a search result from the sorted results causes the result to be added to the first set of data and causes the results whose links were not clicked by the user but that are ranked higher than the clicked result to be automatically added to the second set of data, the first and second sets of data persisted to a computer-readable storage medium</u>.

14.   (Previously Presented) The system of claim 13, the filter component parses the results as a function of one or more of a document property, a context parameter, and a configuration associated with the entry point.

15.   (Original) The system of claim 13, the filter component trained to differentiate between a relevant and a non-relevant result *via* the training data.

16.   (Previously Presented) The method of claim 13, the second set of data categorized as non-relevant comprising random data unrelated to the search context of the user for the entry point.

17.   (Previously Presented) The system of claim 13, the filter component employs statistical analysis to determine whether a result is relevant or non-relevant to the entry point.

18.   (Previously Presented) The system of claim 13, the ranking component employs a technique to determine the degree of relevance of the query results with respect to the relevant data set and the non-relevant data set.

19.   (Original) The system of claim 18, the technique comprising one of a similarity measure and a confidence interval.

20.   (Original) The system of claim 13, the ranking order comprising one of ascending and descending, from the most relevant result to the least relevant result.

21.   (Previously Presented) The system of claim 18, the ranking performed on the relevant query results, the non-relevant results are discarded.

22.   (Currently Amended) A method to filter and rank general-purpose search engine results based on criteria associated with an entry point, comprising:

   executing a query search with the general-purpose search engine accessed through a link associated with the entry point;

   filtering the general-purpose search engine results by tuning the general-purpose search engine based on a set of training data associated with the entry point employed to access the general purpose search engine; and

   ranking the filtered general-purpose search engine results;

   automatically storing a first query result selected by a user in a first data set categorized as relevant;

   automatically storing at least one ~~unselected~~ non-selected query result that is ranked higher than the ~~selected~~ first query result in a second data set categorized as non-relevant upon selection of the first query result; and

   including the first data set and second data set in the set of training data associated with the entry point employed to access the general purpose search engine.

23.   (Original) The method of claim 22, further comprising employing a statistical hypothesis to determine whether a result is relevant or non-relevant to a search context of the entry point.

24.   (Previously Presented) The method of claim 23, the statistical hypothesis employing a threshold in connection with a probability distribution for relevant data and a probability distribution for non-relevant data, respective word probabilities are generated for the search query results and compared to the threshold, the probability distribution for relevant data and the probability distribution for non-relevant data to determine whether the results are relevant or non-relevant.

25.   (Original) The method of claim 24, the threshold employed to bias the decision to mitigate one of a result being deemed non-relevant when the result is relevant and a result being deemed relevant when the result is non-relevant.

26.   (Original) The method of claim 22, further employing a probability distribution analysis or machine learning in connection with the filtering and ranking, wherein suitable probability distributions include a Bernoulli, a binomial, a Pascal, a Poisson, an arcsine, a beta, a Cauchy, a chi-square with N degrees of freedom, an Erlang, a uniform, an exponential, a gamma, a Gaussian-univariate, a Gaussian-bivariate, a Laplace, a log-normal, a rice, a Weibull and a Rayleigh distribution, and the machine learning can classify based on one or more of a word occurrence, a distribution, a page layout, an inlink, and an outlink.

27.   (Original) The method of claim 22, further comprising employing a statistical analysis to rank search query results.

28.   (Original) The method of claim 27, the ranking comprising one of generating word probabilities and employing a confidence interval to determine relevance, and generating a similarity measure comprising one of a cosine distance, the Jaccard coefficient, an entropy-based measure, a divergence measure and/or a relative separation measure to determine similarity.

29.   (Currently Amended) A method to customize a general-purpose search engine to improve context search query results, comprising:

tuning a general-purpose search engine for an entry point by employing a method further comprising:

providing a first set of data categorized as relevant that is used by a component to discern query results relevant to a search context of a user employing the entry point, the entry point provides a link employed to access the general-purpose search engine;

providing a second set of data categorized as non-relevant that is used by the component to discern query results unrelated to the search context, the first set of data and the second set of data are manually provided; ~~and~~

determining whether a query result is relevant or non-relevant to the search context based on the first set of relevant data and the second set of non-relevant data, each query result is compared with both the first set of data and second set of data to determine the relevance of the query result[[.]];

executing a search query with the general purpose search engine to obtain a ranked list of query results;

selecting a link associated with a query result from the list;

automatically adding the selected query result to the first set of data; and

automatically adding non-selected results from the list that are ranked higher than the selected query result to the second set of data upon selection of the selected query result.

30.   (Previously Presented) The method of claim 29, the first set of data categorized as relevant comprising data associated with the search context of the user for the entry point.

31.   (Previously Presented) The method of claim 29, the second set data categorized as non-relevant comprising random data unrelated to the search context of the user for the entry point.

32.   (Original) The method of claim 29, further comprising providing information to associate respective query results with the entry point.

33.   (Previously Presented) The method of claim 29, the first set of data categorized as relevant and the second set of data categorized as non-relevant employed to train the component to learn the features that differentiate relevant data from non-relevant data.

34.   (Currently Amended) A method to automatically customize a general-purpose search engine for an entry point, comprising:

        identifying the entry point;

        executing a query search *via* the entry point that includes a link employed to route to the general-purpose search engine;

        recording a <u>first</u> query result <u>from a ranked list of query results returned from the executed query</u> ~~selected by a user employing the entry point~~ as relevant <u>when a user views the document associated with the first query result</u>;

        recording at least ~~one higher ranked but unselected~~ <u>one second</u> query result <u>whose associated document was not viewed by the user but that is ranked higher than the first query result</u> as non-relevant when ~~a lower ranked~~ <u>the first</u> result is selected <u>for viewing</u> by the user; and

        providing the recorded results to automatically train the filter for the entry point, in order to discriminate between results relevant to a search context of the user for the entry point and results non-relevant to the search context.

35.   (Original) The method of claim 34, the set of relevant data comprising data associated with the search context of the user for the entry point.

36.   (Original) The method of claim 34, the set of non-relevant data comprising data unrelated to the search context of the user for the entry point.

37.   (Original) The method of claim 34, further comprising providing information to associate respective query results with the entry point.

38.   (Original) The method of claim 34, the set of relevant data and the set of non-relevant data employed to train the component to learn the features that differentiate relevant data from non-relevant data.

39.   (Previously Presented) The method of claim 34, the query results selected *via* a click thru technique employing a mouse to select a link associated with the query result by clicking on the link.

40.   (Original) The method of claim 34, further comprising generating a word probability distribution for the relevant recorded results and a word probability distribution for the non-relevant recorded results.

41.   (Cancelled).

42.   (Currently Amended) A computer readable storage medium storing computer executable components that tunes a general-purpose search engine to improve context search query results, comprising:

        a component that receives search query results of a general-purpose search engine and filters the results based on training data sets associated with the search context of a user depending on the entry point that provides a link utilized to arrive at the general-purpose search engine, the training data sets include at least a first category of data explicitly defined to be relevant to the search context and a second category of data explicitly defined to be non-relevant to the search context; and

        a component that ranks the filtered general-purpose search engine results according to the similarity of the search engine results to the training data sets, ~~each search result is compared with both the first category of data and the second category of data~~ wherein selecting a link associated with a first search result from the ranked results causes the first result to be added to the first set of data and causes results that are ranked higher than the first result and have not been selected by the user to be automatically added to the second set of data.

43.   (Currently Amended) A system that receives, filters and ranks general-purpose search engine results, comprising:

means for filtering general-purpose search engine results by determining whether a query result is relevant to a search context of a group of users, the search context is associated with an entry point that includes a link employed to navigate to the general-purpose search engine, the search context further having an associated first set of training data categorized as relevant to the context and an associated second set of training data categorized as non-relevant to the context; and

means for ranking the filtered general-purpose search engine results based on a relevance of the general-purpose search engine results to the search context of the group of users and the entry point as determined by a comparison of the search engine results with the first and second sets of training data, wherein a user viewing a document associated with a first search result from the ranked results causes the first result to be added to the first set of training data and causes the results that are unviewed but ranked higher than the first result to be automatically added to the second set of training data, the first and second sets of training data stored on a computer-readable storage medium.